Molecular Communications: Model-Based and Data-Driven Receiver Design and Optimization

Xuewen Qian, Marco Di Renzo, Andrew W. Eckford

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier

Molecular Communications: Model-Based and Data-Driven Receiver Design and Optimization

XUEWEN QIAN[1], MARCO DI RENZO[1], (SENIOR MEMBER, IEEE) AND ANDREW ECKFORD[2], (SENIOR MEMBER, IEEE)

[1]Laboratoire des Signaux et Systèmes, CentraleSupélec - CNRS - Université Paris Sud, Université Paris-Saclay, Plateau de Saclay, 3 rue Joliot-Curie, 91192, Gif-sur-Yvette (Paris), France (e-mail: qian.xuewen, marco.direnzo@l2s.centralesupelec.fr)

[2] Department of Electrical Engineering and Computer Science, York University, Toronto, Canada (e-mail: aeckford@yorku.ca)

[*]This paper was presented in part at 2018 IEEE ISWCS.

Corresponding author: Marco Di Renzo (e-mail: marco.direnzo@l2s.centralesupelec.fr).

ABSTRACT In this paper, we consider a molecular communication system that is made of a 3D unbounded diffusion channel model without flow, a point transmitter, and a spherical absorbing receiver. In particular, we study the impact of inter-symbol interference, and analyze the performance of different threshold-based receiver schemes. The aim of this paper is to analyze and optimize the receivers by using the conventional model-based approach, which relies on an accurate model of the system, and the emerging data-driven approach, which, on the other hand, does not need any apriori information about the system model and exploits deep learning tools. We develop a general analytical framework for analyzing the performance of threshold-based receiver schemes, which are suitable to optimize the detection threshold. In addition, we show that data-driven receiver designs yield the same performance as receivers that have perfect knowledge of the underlaying channel model.

INDEX TERMS Molecular communications, error probability, receiver design, artificial neural networks.

I. INTRODUCTION

Traditional electromagnetic-based transmission techniques may not be appropriate to enable the communication among nano-devices [1]. Molecular Communications (MC) are, on the other hand, a more suitable and emerging option [2]. In a MC system, the information is transmitted via the release of information particles [2]. If the information is encoded onto the number of particles that are released, the corresponding modulation scheme is referred to as Concentration Shift Keying (CSK) modulation.

In MC systems, diffusion [3] is the easiest option to enable information particles propagate from the transmitter to the receiver. Due to the intrinsic characteristics of diffusion, the resulting transmission channel is usually affected by non-negligible Inter-Symbol Interference (ISI) which, if not taken into account for system optimization, may severely degrade the system performance [4–8]. The enzyme-based MC system [9] is one of the available schemes to mitigate the intrinsic ISI in MC systems. If the data rate is high, however, the ISI may not be negligible, and the approach in [9] may not provide satisfactory performance. For this reason, we focus our attention on optimizing MC systems in the presence of ISI. Developing solutions to reduce the impact of ISI is an important research topic in MC systems. For example, approaches based on modulation [10], channel coding [11], and receiver design [12] are available in the literature. In the present paper, we focus our attention on developing robust receiver schemes.

In MC systems, a simple approach [9] to demodulate the, e.g., binary symbol is to compare the number of received particles r_i with a fixed threshold τ: if $r_i \leq \tau$, the symbol is detected as 1, otherwise it is detected as 0. The threshold of this threshold-based detector is relatively simple to be optimized in the absence of ISI or if the ISI is negligible. In general, on the other hand, the threshold needs to be optimized by taking the ISI into account in order to minimize the error probability and obtain good communication performance. In [13], the authors have proposed a scheme that uses the number of particles received in the previous time-slot, i.e., r_{i-1} , as the detection threshold in a given time-slot. In [12], the authors have designed an adaptive receiver that combines a channel estimator and a decision-feedback equalizer. The

channel estimator updates the channel parameters and detects the symbols constantly. Further results are available in [14]. Therein, the authors propose a new decoder that divides each slot into sub-slots. According to the number of received particles in each sub-slots, an associated decision rule is adopted and the whole scheme improves the detection performance. Similar results are available in [8].

The aforementioned approaches rely on the knowledge of the channel and system models. This, however, may not always be possible either due to the complexity of modeling the entire system in an accurate manner or due to the complexity of optimizing the resulting system model. These issues can be solved by using Machine Learning (ML) methods. With the help of ML, several schemes have been proposed [15, 16], e.g., for application to Orthogonal Frequency-Division Multiplexing (OFDM) [17], to circumvent these issues. In MC systems, ML-based schemes have been proposed in, e.g., [18, 19], to address the issue of accurate system modeling. Furthermore, the authors of [20] have recently proposed a sequence detection scheme based on sliding bidirectional recurrent neural networks that does not need channel information. Compared with existing ISI mitigation schemes, with the exception of the enzyme-based approach, ML-based schemes are less complex and easier to implement. With the aid of deep learning methods, the authors of [20] have shown that their proposed scheme is capable of automatically learning the whole system from empirical data and of performing data detection without using complex channel estimation and data equalization techniques.

In this paper, motivated by the promising results obtained in [20], we study the possibility of optimizing the receiver design of MC systems in the presence of ISI by using Artificial Neural Networks (ANNs). In particular, our implementation is based on feed-forward ANNs with fully-connected layers. Our study shows that ANNs without prior knowledge of the system model are capable of providing the same performance as conventional detection methods that rely on the perfect knowledge of the underlaying system model. In particular, the novelty and contribution of this paper can be summarized as follows.

- We compute the Bit Error Rate (BER) of many threshold-based detection schemes. Compared with other frameworks, the proposed approach takes the background noise and the ISI into account in an accurate manner. In particular, we show that the proposed ANN-based implementation is capable of estimating the optimal threshold that minimizes the BER, as opposed to sub-optimal solutions where the dynamic nature of the ISI is not taken into account. Our approach is, therefore, particularity suitable for optimal threshold-based receiver implementations without prior knowledge of the system model.
- We develop different receiver schemes, both based on conventional detection theory and by applying recent results on data-driven optimization based on ANNs. In particular, we consider receiver structures that account for one-bit, two-bit, or all-bit (genie-aided) prior knowledge. We show that both model-based and data-driven approaches yield similar performance, with the latter approach having the benefits of not requiring any a priori information about the system model.

Compared with the companion conference version [21], this paper is largely expanded, since it encompasses the modeling, analysis, and optimization of different types of receivers whose major difference is the number of previously-detected bits that they use for demodulation. In [21], in fact, only the zero-bit memory receiver was considered. In [21], in addition, no general framework to compute the BER was proposed. The approach proposed in this paper, on the other hand, can be applied to receivers with an arbitrary number of past bits used for detection. The corresponding architectures of the ANNs are proposed as well.

The remainder of this paper is organized as follows. In Section II, the system model is introduced. In Section III, model-based detection schemes are proposed. In Section IV, data-driven detection schemes are developed and optimized. In Section V, the proposed schemes are validated via numerical simulations and illustrations are provided. Finally, Section VI concludes this paper.

II. SYSTEM MODEL

Figure 1 depicts the main components of a MC system. The transmitter generates the information particles, which are released into the channel. The transmitter is assumed to be small enough to be considered as a point. We assume that the information particles diffuse randomly and independently of each other through the medium (Brownian motion). Even though a large number of information particles are emitted, not all of them reach the receiver in the considered time-slot. Some information particles remain in the channel and reach the receiver in subsequent time-slots: this causes the ISI. If not appropriately taken into account, the ISI may severely degrade the performance of MC systems. As an example, we consider a spherical absorbing-type receiver [22, 23].

We assume that the temperature is constant and the viscosity η remains unchanged during the whole transmission duration. The diffusion coefficient D [1], thus, remains constant as well. In the considered system model, no extra energy is needed since particles diffuse freely.

We consider a 3D unbounded diffusion channel model without flow, as illustrated in Fig. 2. By assuming the transmitter located at $\boldsymbol{a} = (0,0,0)$ and the receiver at $\boldsymbol{b} = (b_x, b_y, b_z)$, the hitting rate of each information particle can be expressed as follows [13, 23]:

$$f_{hit}^{3D}(t) = \frac{r(d-r)}{d\sqrt{4\pi D t^3}} e^{-\frac{(d-r)^2}{4Dt}} \tag{1}$$

where $\|\boldsymbol{a}-\boldsymbol{b}\| = d$ is the distance between the transmitter and the center of the receiver, and r is the radius of the receiver that is assumed to have a spherical shape.

For ease of illustration, an On-Off Keying (OOK) modulation scheme is considered. At the ith slot, the transmitter

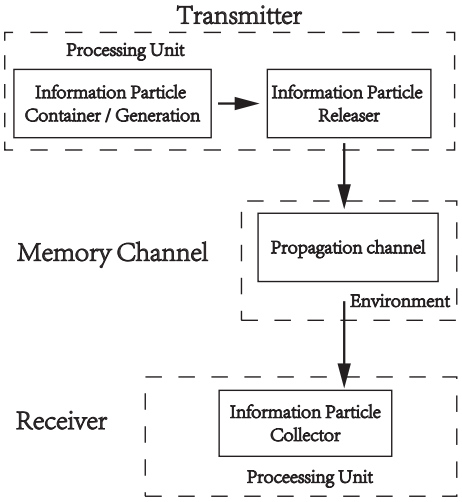

FIGURE 1: Block diagram of a typical MC system.

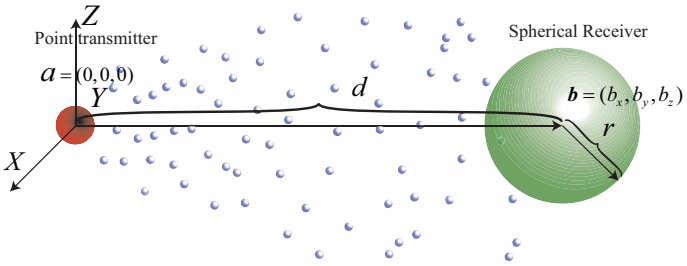

FIGURE 2: The 3D unbounded molecular channel model without flow including a point transmitter and a spherical absorbing receiver.

releases N_{TX} information particles into the environment when the symbol is $s_i = 1$, otherwise the transmitter does not release any particles. We assume that the transmitter can release the N_{TX} information particles in a very short time so that the release time effect of the transmitter on the received signal is negligible.

The hitting probability of an absorbing receiver, i.e., the probability to absorb one particle after t seconds that the information particle is released, can be expressed as follows:

$$P_{hit}(t) = \int_0^t f_{hit}(t)dt \tag{2}$$

From (1) and (2), we have:

$$P_{hit}(t) = \int_0^t f_{hit}(t)dt = \frac{r}{d} erfc(\frac{d-r}{\sqrt{4Dt}}) \tag{3}$$

where: $erf(y) = \int_0^y \frac{2}{\sqrt{\pi}} e^{-x^2} dx$ and $erfc(y) = 1 - erf(y)$.

Therefore, during the $(i-1)$th time slot after releasing the particle, the probability that one particle hits the receiver is:

$$P_{i-1} = \int_{(i-1)T}^{iT} f_{hit}(t)dt \tag{4}$$

TABLE 1: Simulation parameters

Parameter	Value
λ_0	$100s^{-1}$
Receiver radius r	45 nm
Distance d	500 nm
Diffusion coefficient D	$4.265 * 10^{-10} m^2/s$
Discrete time length ΔT	9 us
Slot length T	$30\Delta T$
Channel length L	5

Then, we obtain:

$$P_{i-1} = \frac{r}{d}\{erfc(\frac{d-r}{\sqrt{4DiT}}) - erfc(\frac{d-r}{\sqrt{4D(i-1)T}})\} \tag{5}$$

Let $C_j = N_{TX} P_j$ denote the average received particles at the jth time-slot if N_{TX} particles are released. Then, the number of received particles [14] at the ith time-slot follows the Poisson distribution as follows:

$$r_i \sim Poisson(I_i + s_i C_0) \tag{6}$$

where $I_i = \lambda_0 T + \sum_{j=1}^{\infty} s_{i-j} C_j$ is the sum of ISI and background noise, and λ_0 is the background noise power per unit time.

More precisely, the probability of receiving r_i information particles is:

$$P(r_i | I_i + s_i C_0) = \frac{e^{-(I_i + s_i C_0)} (I_i + s_i C_0)^{r_i}}{r_i!} \tag{7}$$

For ease of tractability, we assume that C_i for $i > L$ is small enough to be integrated into the background noise, and we denote by L the length of the Poisson channel. Therefore, C_i for $1 \leq i \leq L$ contributes, in practice, to the ISI. The signal-to-noise ratio (SNR) can be defined as follows:

$$SNR = 10 log_{10} \frac{C_0}{2\lambda_0 T} \tag{8}$$

where the information bits are assumed to be equiprobable.

Accordingly, given a certain SNR value, the number of released particles, N_{TX}, is:

$$N_{TX} = 2\lambda_0 T 10^{\frac{SNR}{10}} / P_0 \tag{9}$$

For future reference, the system parameters of a typical MC system are listed in Table 1.

III. MODEL-BASED RECEIVERS DESIGN IN MOLECULAR COMMUNICATIONS

In this section, we study some receivers in the presence of ISI with the objective of computing their bit error rate (BER) performance and optimizing their parameters in order to minimize the BER. For all cases, the system model is the same as in the previous section. We consider different types of threshold-based detectors, whose main difference consists of the amount of prior information, i.e., the number of previous bits that they use for demodulation.

A. OPTIMAL ZERO-BIT MEMORY RECEIVER

We study a threshold-based zero-bit memory receiver. The demodulation threshold is denoted by τ. Let $\bar{s}_i$ be the estimate of symbol s_i at time-slot i. The demodulation rule can be formulated as follows:

$$\bar{s}_i = \begin{cases} 0, & r_i \leq \tau \\ 1, & r_i > \tau \end{cases} \tag{10}$$

The traditional approach to determine the threshold τ is obtained by imposing $P(r_i = \tau | s_i = 0) = P(r_i = \tau | s_i = 1)$. The rationale behind this approach is that the values of C_i for $1 \leq i \leq L$ are unknown, but the averaged ISI, equal to $\sum_{i=1}^{L} C_i/2$, is known. Under these assumptions, the probability of receiving r_i particles conditioned upon s_i can be written as follows:

$$P_{app}(r_i|s_i) = \frac{e^{-\lambda|_{s_i}} (\lambda|_{s_i})^{r_i}}{r_i!} \tag{11}$$

where $\lambda|_{s_i} = C_0 s_i + \frac{\sum_{j=1}^{L} C_j}{2} + \lambda_0 T$.

By imposing the equality $P_{app}(r_i|s_i = 0) = P_{app}(r_i|s_i = 1)$, we obtain:

$$\tau = \frac{C_0}{ln(1 + \frac{C_0}{\sum_{i=1}^{L} C_i/2 + \lambda_0 T})} \tag{12}$$

This approach is, however, sub-optimal. If, in fact, the slot length is long enough, then the term $\sum_{i=1}^{L} C_i/2$ is a good approximation for the ISI. If the slot length is short, on the other hand, the ISI changes according to the previously transmitted symbols and $\sum_{i=1}^{L} C_i/2$ is not a good approximation anymore. Thus, the obtained demodulation threshold is not the optimal choice anymore.

In the following proposition, we develop the optimal demodulation threshold that minimizes the BER. To this end, we propose also a new accurate analytical formulation of the BER. For ease of notation, we denote by $\boldsymbol{s}_{i-1} = \{s_{i-1}, s_{i-2}, ...s_{i-L}\}$ the vector of bits that are transmitted in the L time-slots preceding the ith time-slot of interest.

Proposition 1. *The optimal threshold that minimizes the BER of the zero-bit memory receiver is as follows:*

$$(\tau^*, P_e^*) = \arg\min_{\tau} \quad P_e(\tau) \tag{13}$$

where $P_e(\tau)$ is the BER as a function of τ:

$$P_e(\tau) = \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} P_e(\boldsymbol{s}_{i-1}, \tau) \tag{14}$$

and:

$$\begin{aligned} P_e(\boldsymbol{s}_{i-1}, \tau) = \frac{1}{2}[&Q(\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j, \lceil \tau \rceil) \\ &+ 1 - Q(\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j + C_0, \lceil \tau \rceil)] \end{aligned} \tag{15}$$

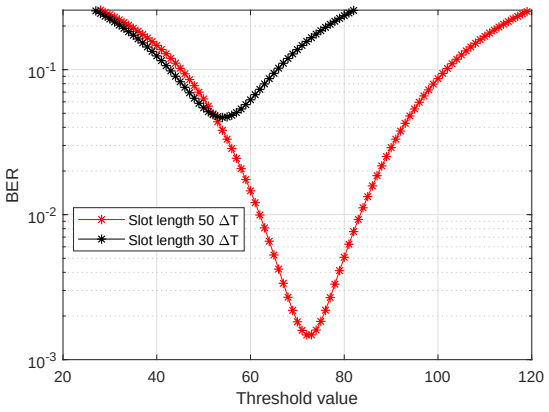

FIGURE 3: BER in (14) as a function of τ (the SNR is 30 dB).

Proof. The BER is defined as follows:

$$\begin{aligned} P_e(\boldsymbol{s}_{i-1}, \tau) = \frac{1}{2}[&P(r_i \geq \tau | s_i = 0, \boldsymbol{s}_{i-1}) \\ &+ P(r_i < \tau | s_i = 1, \boldsymbol{s}_{i-1})] \end{aligned} \tag{16}$$

where:

$$\begin{aligned} P(r_i \geq \tau | s_i = 0, \boldsymbol{s}_{i-1}) &= P(r_i \geq \tau | \lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j) \\ &= \sum_{k=\lceil \tau \rceil}^{\infty} \frac{e^{-(\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j)} (\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j)^k}{k!} \\ &= Q(\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j, \lceil \tau \rceil) \end{aligned} \tag{17}$$

where $Q(\lambda, n) = \sum_{k=n}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!}$ is the incomplete Gamma function and $Q(\lambda, 0) = 1$. Similarly, we have:

$$\begin{aligned} P(r_i < \tau | s_i = 1, \boldsymbol{s}_{i-1}) &= P(r_i < \tau | \lambda_0 T + \sum_{j=0}^{L} s_{i-j} C_j) \\ &= \sum_{k=0}^{\lceil \tau \rceil - 1} \frac{e^{-(\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j + C_0)} (\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j + C_0)^k}{k!} \\ &= 1 - \sum_{k=\lceil \tau \rceil}^{\infty} \frac{e^{-(\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j + C_0)} (\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j + C_0)^k}{k!} \\ &= 1 - Q(\lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j + C_0, \lceil \tau \rceil) \end{aligned} \tag{18}$$

From (17) and (18), we obtain (15). Finally, the BER is obtained by averaging (14) with respect to the vector $\boldsymbol{s}_{i-1}$. □

The optimal detection threshold, τ, is obtained by minimizing the BER (see (13)). In Fig. 3, we depict (14) as a function of τ. We observe that an optimal value of τ exists that minimizes the BER and that it depends on the time slot duration T, i.e., the amount of ISI.

B. OPTIMAL ONE-BIT MEMORY RECEIVER

In this section, we study and optimize the performance of a one-bit memory receiver, which has more prior information than the zero-bit memory receiver. The receiver can be formulated as follows:

$$\bar{s}_i = \begin{cases} 0, & r_i \leq \tau|_{s_{i-1}} \\ 1, & r_i > \tau|_{s_{i-1}} \end{cases} \tag{19}$$

where $\tau|_{s_{i-1}}$ denotes the threshold for the ith symbol when the previously transmitted symbol is s_{i-1}. Since the exact value of s_{i-1} is unknown, the estimate $\bar{s}_{i-1}$ is employed instead, i.e., $\tau|_{\bar{s}_{i-1}}$ is used.

In simple terms, in contrast to the zero-bit memory receiver that accounts only for the number of received particles in the time-slot of interest, the one-bit memory detector adapts the detection threshold as a function of the previously transmitted bit (that is, in practice, replaced by its estimate). Therefore, the detection threshold changes from time-slot to time-slot, but a better approximation of the ISI is obtained.

The following proposition yields the optimal value of the detection threshold, by using the same line of thought as for the zero-bit memory detector.

Proposition 2. *The optimal detection threshold of the one-bit memory receiver can be formulated as follows:*

$$\tau^*|_{s_{i-1}} = \arg\min_{\tau} \quad P_e(\tau, s_{i-1}) \tag{20}$$

where the BER is as follows:

$$P_e(\tau, s_{i-1}) = \frac{1}{2^{L-1}} \sum_{s_{i-2},\cdots,s_{i-L}} P_e(\boldsymbol{s}_{i-1}, \tau) \tag{21}$$

In order to compute the optimal threshold for the ith time-slot, the previously transmitted symbol s_{i-1} is assumed to be known. In practice, this is not possible, since only its estimates is available, as discussed already. Therefore, the BER needs to take this into account. The BER of the one-bit memory receiver is given in the following theorem.

Theorem 1. *The BER of the one-bit memory detector can be formulated as follows:*

$$P_e = \frac{m+n}{2} \tag{22}$$

where m and n are the solutions of the following equations:

$$m = \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} \sum_{\bar{s}_{i-1}} Q(\lambda|_{\boldsymbol{s}_{i-1}, s_i=0}, \lceil \tau|_{\bar{s}_i} \rceil) \Psi(s_{i-1}, \bar{s}_{i-1}, m, n) \tag{23}$$

$$n = \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} \sum_{\bar{s}_{i-1}} (1 - Q(\lambda|_{\boldsymbol{s}_{i-1}, s_i=1}, \lceil \tau|_{\bar{s}_i} \rceil)) \Psi(s_{i-1}, \bar{s}_{i-1}, m, n) \tag{24}$$

where $\tau|_{\bar{s}_{i-1}}$ is the optimal threshold that corresponds to the previously detected bit $\bar{s}_{i-1}$, $\lambda|_{\boldsymbol{s}_{i-1}, s_i=0}$ is the average number of received particles by conditioning on the current symbol being 0 and the previous L symbols being $\boldsymbol{s}_{i-1}$, i.e., $\lambda|_{\boldsymbol{s}_{i-1}, s_i=0} = \sum_{j=1}^{L} C_j s_{i-j} + \lambda_0 T$. Similarly, $\lambda|_{\boldsymbol{s}_{i-1}, s_i=1} = \sum_{j=1}^{L} C_j s_{i-j} + \lambda_0 T + C_0$. The function $\Psi(s_{i-1}, \bar{s}_{i-1}, m, n)$ is defined as follows:

$$\Psi(s_{i-1}, \bar{s}_{i-1}, m, n) = \begin{cases} m, & (s_{i-1}=0, \bar{s}_{i-1}=1) \\ 1-m, & (s_{i-1}=0, \bar{s}_{i-1}=0) \\ n, & (s_{i-1}=1, \bar{s}_{i-1}=0) \\ 1-n, & (s_{i-1}=1, \bar{s}_{i-1}=1) \end{cases} \tag{25}$$

Proof. See Appendix VII-A. □

C. OPTIMAL K-BIT MEMORY RECEIVER

Inspired by the one-bit memory detector in Section III-B, we generalize this receiver design by considering a generic K-bit memory receiver. It is worth nothing that K may be set equal to L, which is the actual length of the ISI channel. This setup yields the optimal performance but needs more a priori information on the previously detected bits, which increases the complexity of the receiver.

The optimal detection threshold and BER are given in the following proposition and theorem, respectively.

Proposition 3. *The optimal detection threshold of the K-bit memory receiver can be formulated as follows:*

$$\tau^*|_{s_{i-1},...,s_{i-K}} = \arg\min_{\tau} \quad P_e(\tau, s_{i-1}, ..., s_{i-K}) \tag{26}$$

where the BER is as follows:

$$P_e(\tau, s_{i-1}, ..., s_{i-K}) = \frac{1}{2^{L-K}} \sum_{s_{i-K+1},\cdots,s_{i-L}} P_e(\boldsymbol{s}_{i-1}, \tau) \tag{27}$$

Since the exact symbols s_{i-j}, for$1 \leq j$ are unknown, the estimates $\bar{s}_{i-1}, ..., \bar{s}_{i-K}$ are used to perform the detection:

$$\bar{s}_i = \begin{cases} 0, & r_i \leq \tau|_{\bar{s}_{i-1},...,\bar{s}_{i-K}} \\ 1, & r_i > \tau|_{\bar{s}_{i-1},...,\bar{s}_{i-K}} \end{cases} \tag{28}$$

Theorem 2. *The BER of the K-bit memory receiver can be approximated by using (22), where m and n are the solution of the equations:*

$$m = \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} \sum_{\bar{s}_{i-1},...,\bar{s}_{i-K}} Q(\lambda|_{\boldsymbol{s}_{i-1}, s_i=0}, \lceil \tau|_{\bar{s}_{i-1},...,\bar{s}_{i-K}} \rceil) \prod_{j=1}^{K} \Psi(s_{i-j}, \bar{s}_{i-j}, m, n)$$

$$n = \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} \sum_{\bar{s}_{i-1},...,\bar{s}_{i-K}} (1 - Q(\lambda|_{\boldsymbol{s}_{i-1}, s_i=1}, \lceil \tau|_{\bar{s}_{i-1},...,\bar{s}_{i-K}} \rceil)) \prod_{j=1}^{K} \Psi(s_{i-j}, \bar{s}_{i-j}, m, n)$$

Proof. See Appendix VII-B. It is worth mentioning that the obtained expression of the BER is an approximation for general values of K. The details of the approximation are available in the appendix. □

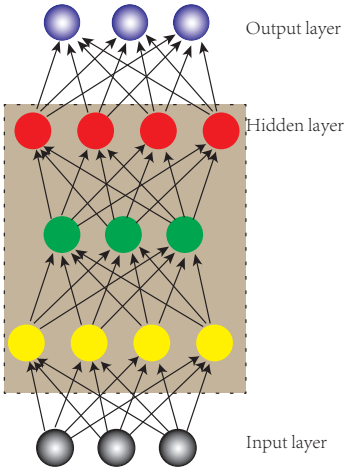

FIGURE 4: Typical structure of a feed-forward ANN with fully-connected layers.

IV. DATA-DRIVEN RECEIVER DESIGN IN MOLECULAR COMMUNICATIONS

In the previous section, we have optimized the operation of receivers by assuming that the underlying system model is perfectly known. In this section, we do not rely on this assumption anymore, and take advantage of feed-forward ANNs with fully-connected layers and deep learning [24] to optimize the design of molecular receivers. The architecture of a typical feed-forward ANN with fully-connected layers is depicted in Fig 4, and it consists of an input layer, several hidden layers, and an output layer. The nodes of the hidden layers are referred to as neurons.

The objective of this section is to describe how to design and optimize receivers by using ANNs and by training an ANN by using only empirical data, i.e., the number of received particles in the presence of ISI. We describe each receiver studied in the previous section in the following sub-sections.

A. DATA-DRIVEN DESIGN OF ZERO-BIT MEMORY RECEIVER

The objective of an ANN-based design is to identify an ANN structure that demodulates the transmitted data by minimizing the BER. An ANN-based zero-bit memory demodulator is a system whose input consists of the received information particles r_i at the ith time-slot, and the outputs are the probabilities that the transmitted bit is 0 or 1, i.e., $P_i(s_i = 0|r_i)$ and $P_i(s_i = 1|r_i)$, respectively. Since, $P_i(s_i = 1|r_i) + P_i(s_i = 0|r_i) = 1$, only one of the two probabilities is needed. In the sequel, we use the notation $P_i = P_i(s_i = 1|r_i)$. Based on the inputs, the ANN demodulate the received data as follows:

$$\bar{s}_i = \begin{cases} 0, & P_i \leq 0.5 \\ 1, & P_i > 0.5 \end{cases} \tag{29}$$

where the threshold 0.5 accounts for the fact that the bits are equiprobable.

In order to train the ANN, we consider a supervised learning approach, i.e., we compute the parameters (e.g., the bias factors and the weights) of the ANN by using a known sequence of transmitted bits. In particular, we use the Bayesian regularization back propagation technique, which updates the weights and biases by using the Levenberg-Marquardt optimization algorithm. A hyperbolic tangent sigmoid activation function is employed in all neurons. The set of parameters that are used to train and operate the ANN are the following: The number of hidden layers is 10, the number of neurons per layer is 5, the learning rate is 0.01, the training epoch is 200, the total number of training bits is 50000, and the number of test bits is 100000. In particular, the training is performed in a batch mode, and the number of bits in each batch is 1000. This setup is used to obtain the numerical results in the next section.

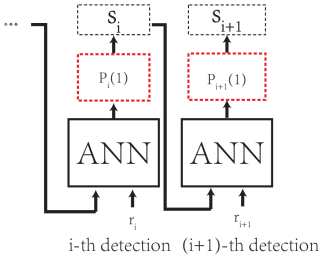

FIGURE 5: Data-driven one-bit memory receiver.

B. DATA-DRIVEN DESIGN OF ONE-BIT MEMORY RECEIVER

If the one-bit memory receiver is considered, the input of the ANN is not just the number of received particles at the ith time-slot, r_i, but also the estimated symbol at the $(i-1)$th time-slot, $\bar{s}_{i-1}$. In mathematical terms, the output estimate of the ANN can be formulated as follows:

$$\bar{s}_i = \begin{cases} 0, & P(s_i = 1|r_i, \bar{s}_{i-1}) \leq 0.5 \\ 1, & P(s_i = 1|r_i, \bar{s}_{i-1}) > 0.5 \end{cases}$$

A block diagram representation of the ANN-based architecture is depicted in Fig. 5.

As far as the ANN architecture is concerned, the same system setup as for the zero-bit memory receiver is considered with the only exception that the number of hidden layers is equal to 5 and the number of neurons per layer is 4.

C. DATA-DRIVEN DESIGN OF K-BIT MEMORY RECEIVER

By using the same line of thought as for the one-bit memory receiver, the decision rule of the K-bit memory receiver can be formulated as follows:

$$\bar{s}_i = \begin{cases} 0, & P(s_i = 1|r_i, \bar{s}_{i-1}, ..., \bar{s}_{i-K}) \leq 0.5 \\ 1, & P(s_i = 1|r_i, \bar{s}_{i-1}, ..., \bar{s}_{i-K}) > 0.5 \end{cases}$$

The corresponding block diagram is illustrated in Fig. 6.

In this case, in particular, the training data-unit is constituted by the vector $\{r_i, s_{i-1}, ...s_{i-K}; s_i\}$. The same system setup as for the one-bit memory receiver is considered to obtain the numerical results illustrated in the next section.

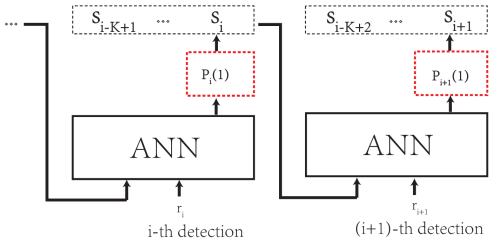

FIGURE 6: Data-driven K-bit memory receiver.

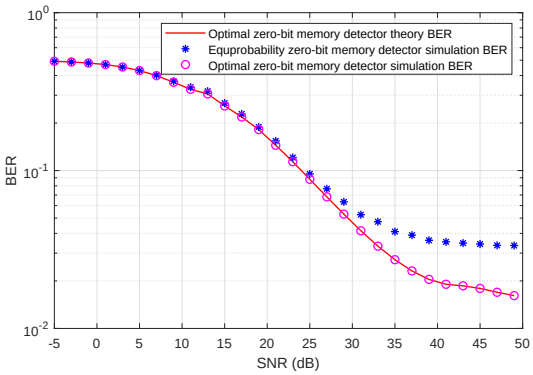

FIGURE 7: BER of optimal vs. conventional zero-bit memory receiver - $T = 30\Delta T$.

V. NUMERICAL RESULTS

In this section, we report and describe some simulation results in order to validate the analysis, design, and optimization of the proposed receivers for application to molecular communications. In addition, we compare both model-based and data-driven designs.

As far as Monte Carlo simulations are concerned, the MC system is assumed to be perfectly synchronized. Accordingly, the hitting rate at each ΔT can be obtained directly from (1), and the number of received particles can be, thus, com-puted from (6) without the need of implementing particle-based Monte Carlo simulations. This approach reduces the simulation time without compromising, under the considered system model, the accuracy of the results.

A. ZERO-BIT MEMORY RECEIVERS

In Fig. 7, we observe that the proposed design based on optimizing the detection threshold that minimizes the BER provides us with better performance than the sub-optimal design. We note, in particular, that for each SNR the optimal detection threshold is used. We observe, in addition, a very good accuracy of the proposed analytical framework. Notably, Fig. 8 and Fig. 9 provide us with a simple representation of the receiver sub-optimality discussed in Sec. III-A. First of all, we observe that the theoretical and empirical distributions of the received number of particles are different. More importantly, we observe that the empirical distributions cross each other in correspondence of the estimated optimal detection threshold, while the approximated ones cross each other in a different point. This justifies the reason why our approach yields the optimum and a better BER.

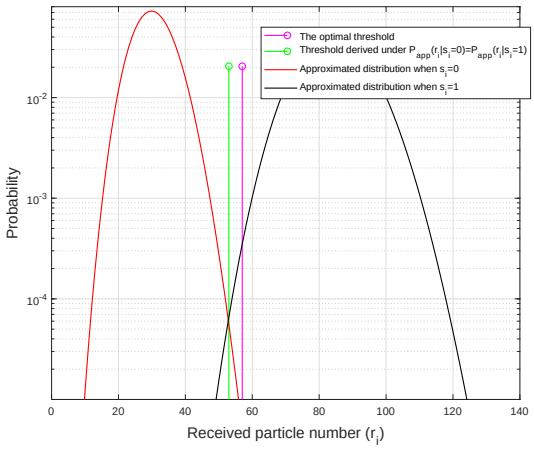

FIGURE 8: Approximated distributions of the received bits from (11) (the SNR is 25 dB).

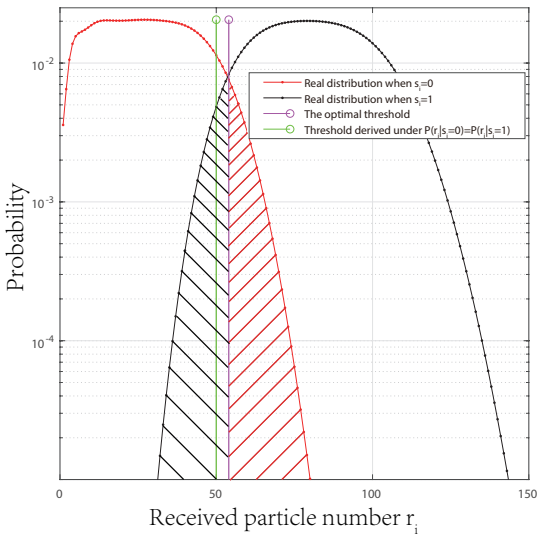

FIGURE 9: Empirical distributions of the received bits (the SNR is 25 dB).

The results shown in Fig. 8 and Fig. 9 are, therefore, very important in order to highlight the sub-optimality of the demodulation thresholds that have been used in the literature to date. The results in Fig. 7, in addition, highlight the advantages of the proposed optimal thresholds in the context of MC systems design and optimization.

In Fig. 10, we compare the BER of the ANN-based demodulator against the model-based receiver design. We observe a good accuracy, which confirms the correct optimization of the ANN, and, at the same time, the correct calculation of the analytical framework. In Fig. 11, we compare the optimal threshold computed numerically from (13) as a function of the SNR, and the demodulation threshold that is learned by

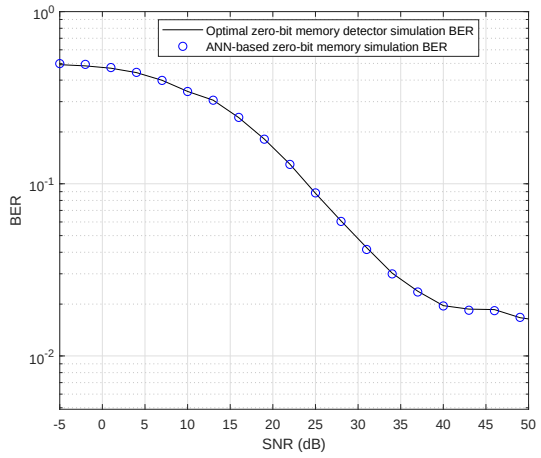

FIGURE 10: The ANN-based receiver achieves the same BER performance as the optimal zero-bit receiver - $T = 30\Delta T$.

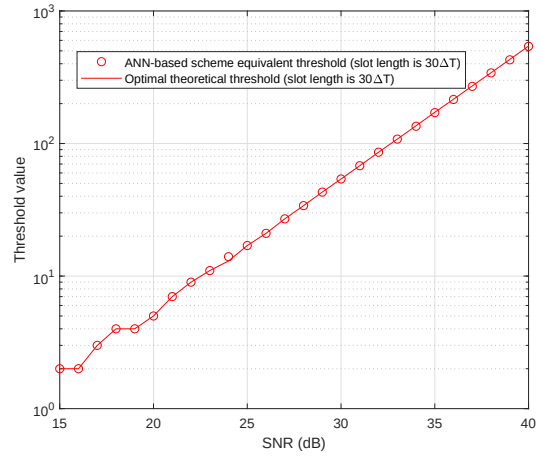

FIGURE 11: Optimal detection thresholds of model-based and data-driven schemes. The ANN-based receiver auto-matically emulates the optimal zero-bit memory receiver. -$T = 30\Delta T$.

the ANN-based demodulator. In the latter case, the threshold is obtained, after completing the training of the ANN, and identifying the input, i.e., the number of information particles, for which the output probability is equal to 0.5. We observe that the ANN-based implementation is capable of learning the demodulation threshold in a very accurate manner. This result is very interesting, as it allows us to unveil the hidden behavior of the optimized ANN. It highlights, in particular, that the optimized ANN is, indeed, a threshold-based demodulator.

In order to further test the robustness of our ANN-based design, we consider another case study where the distance between the transmitter and receiver is short [25]. In this case, the hitting rate can be formulated as follows:

$$f_{short}(t) = \frac{1}{2}[erf(\tau_1) + erf(\tau_2)] - \frac{\sqrt{Dt}}{d\sqrt{\pi}}[e^{-\tau_1^2} - e^{-\tau_2^2}]$$

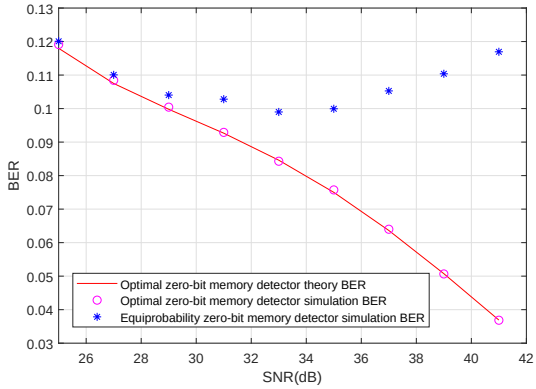

FIGURE 12: System model assuming a short distance between transmitter and receiver. BER of the optimal vs. sub-optimal (i.e., based on the sub-optimal threshold) zero-bit memory receiver - $T = 30\Delta T$.

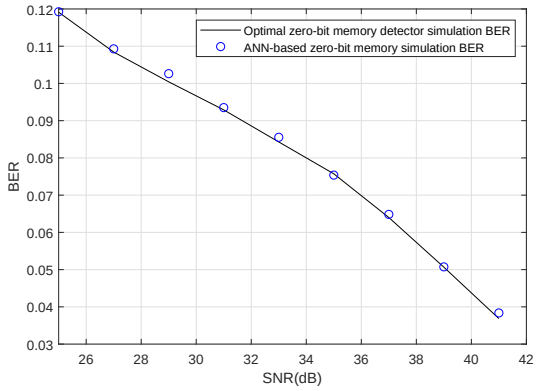

FIGURE 13: System model assuming a short distance between transmitter and receiver. The data-driven receiver achieves the same BER performance as the optimal (i.e., based on the optimal threshold) zero-bit receiver - $T = 30\Delta T$.

where $\tau_1 = \frac{r+d}{\sqrt{4Dt}}$ and $\tau_1 = \frac{r-d}{\sqrt{4Dt}}$.

The rest of the equations can be obtained from this new expression of the hitting rate. As for the simulation parameters, they are the same as those in Table 1, with the exception of the distance $d = 100nm$ and the channel length $L = 2$.

The corresponding results are illustrated in Fig. 12 and Fig. 13. We observe that similar performance trends as for the first case study are obtained. There exists a gap between the BER of optimal and sub-optimal zero-bit receivers. In the following sections, therefore, we will consider only the channel where the distance between the transmitter and the receiver is large.

B. ONE-BIT MEMORY RECEIVERS

In Fig. 14, we compare the BER of the optimal and sub-optimal one-bit threshold receivers. Also in this case, we observe that better performance is obtained by using the proposed optimal design. In addition, the numerical results confirm the correctness of our analytical framework. In Fig.15, we observe that the proposed ANN-based design yields the

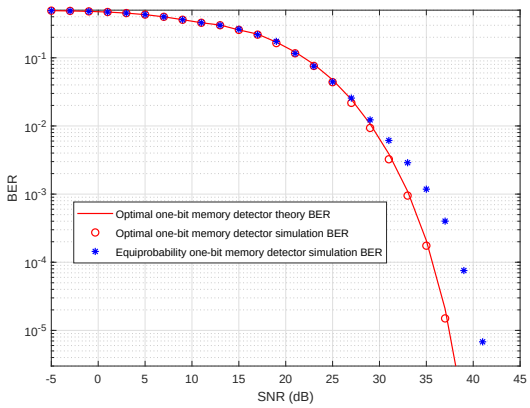

FIGURE 14: BER of optimal vs. sub-optimal one-bit memory receiver. The performance gap between optimal and sub-optimal receiver gets smaller since the ISI is modeled more accurately - $T = 30\Delta T$.

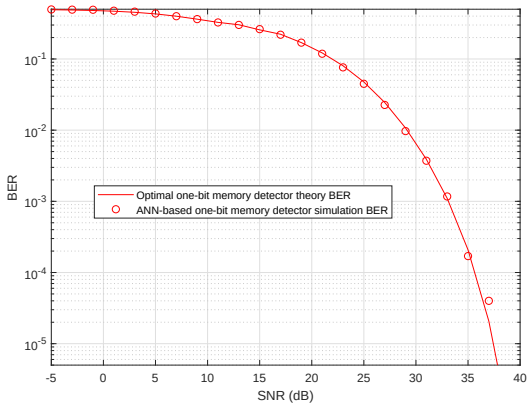

FIGURE 15: The ANN-based receiver achieves the same BER performance as the optimal one-bit receiver - $T = 30\Delta T$.

same results as the model-based approach, which, however, assumes perfect knowledge of the system model.

C. K-BIT MEMORY RECEIVERS

In this section, finally, we consider the design of receivers that exploit more than one bit for improving the performance.

In Fig. 16, we compare the BER of the optimal and sub-optimal receivers by assuming $K = 2$. We observe that, in this case, the sub-optimal receiver is closer to the optimal one, if compared to the case studies with $K = 0$ and $K = 1$. We have observed, in general, that the larger the number of bits of memory is, the closer the BER of optimal and sub-optimal receivers are.

In Fig. 17 and Fig. 18, we compare model-based and ANN-based receiver designs, and we observe a good agreement. In particular, Fig. 18 highlights the improved performance that is obtained by increasing the number of bits of memory, which, for the considered setup, is $K = L$, i.e., the actual length of the ISI.

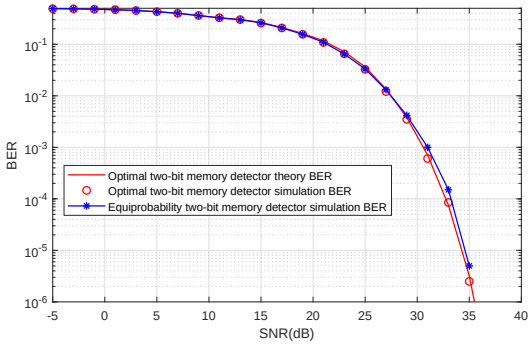

FIGURE 16: Two-bit memory detector: Comparison between the optimal and sub-optimal setups of the demodulation thresholds. The more the number of memory bits, the better the ISI is modeled. As the memory length approaches the channel length, thus, the optimal threshold converges towards the conventional threshold - $T = 30\Delta T$.

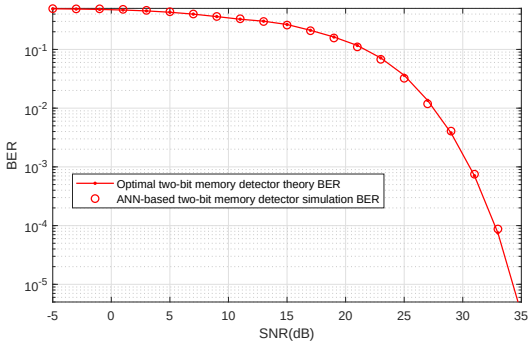

FIGURE 17: The ANN-based two-bit memory receiver achieves the same BER performance as the optimal two-bit receiver - $T = 30\Delta T$.

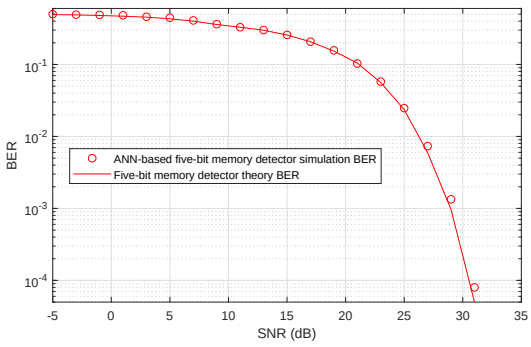

FIGURE 18: The BER performance of ANN-based and model-based L-bit memory detectors overlap with each other - $T = 30\Delta T$.

VI. CONCLUSION

In this paper, we have introduced a new analytical framework to compute the BER of a MC system that uses threshold-based demodulators. By modeling the receiver as an ANN, in addition, we have proved that data-driven receivers provide similar performance as those that are optimized based on the exact knowledge of the channel model. From the considered

ANN-based receiver design, in addition, we have shown that the resulting ANN architecture results in a threshold-based receiver whose threshold coincides with that predicted theoretically. This is an interesting result for better optimizing and further understanding MC systems.

VII. APPENDIX

A. PROOF OF THE BER OF THE ONE-BIT MEMORY RECEIVER

The BER is defined as follows:

$$P_e = \frac{1}{2}\left[P(\bar{s}_i = 1|s_i = 0) + P(\bar{s}_i = 0|s_i = 1)\right] \quad (30)$$

where $P(\bar{s}_i = 1|s_i = 0)$ is the probability of detecting s_i equal to 1 when the transmitted symbol is 0. We have the following:

$$\begin{aligned}
&P(\bar{s}_i = 1|s_i = 0)\\
=& \sum_{\boldsymbol{s}_{i-1},\bar{\boldsymbol{s}}_{i-1}} P(\bar{s}_i = 1|s_i = 0, \boldsymbol{s}_{i-1}, \bar{\boldsymbol{s}}_{i-1})\\
&P(\bar{\boldsymbol{s}}_{i-1}, \boldsymbol{s}_{i-1}, ..., s_{i-L})\\
=& \sum_{\boldsymbol{s}_{i-1},\bar{\boldsymbol{s}}_{i-1}} P(r_i \geq \tau|_{\bar{\boldsymbol{s}}_{i-1}}|s_i = 0, \boldsymbol{s}_{i-1}, \bar{\boldsymbol{s}}_{i-1})\\
&P(\bar{\boldsymbol{s}}_{i-1}|\boldsymbol{s}_{i-1})P(\boldsymbol{s}_{i-1})P(\boldsymbol{s}_{i-2}, ..., s_{i-L})\\
=& \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1},\bar{\boldsymbol{s}}_{i-1}} P(\bar{\boldsymbol{s}}_{i-1}|\boldsymbol{s}_{i-1}) Q(\lambda|_{\boldsymbol{s}_{i-1}, s_i=0}, \lceil \tau|_{\bar{\boldsymbol{s}}_{i-1}} \rceil)
\end{aligned} \quad (31)$$

where $\lambda|_{\boldsymbol{s}_{i-1}, s_i=0} = \lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j$. By using similar steps, we obtain:

$$\begin{aligned}
&P(\bar{s}_i = 0|s_i = 1)\\
=& \sum_{\boldsymbol{s}_{i-1},\bar{\boldsymbol{s}}_{i-1}} P(\bar{s}_i = 0|s_i = 1, \boldsymbol{s}_{i-1}, \bar{\boldsymbol{s}}_{i-1})\\
&P(\bar{\boldsymbol{s}}_{i-1}, \boldsymbol{s}_{i-1}, ..., s_{i-L})\\
=& \sum_{\boldsymbol{s}_{i-1},\bar{\boldsymbol{s}}_{i-1}} P(r_i < \tau|_{\bar{\boldsymbol{s}}_{i-1}}|s_i = 1, \boldsymbol{s}_{i-1}, \bar{\boldsymbol{s}}_{i-1})\\
&P(\bar{\boldsymbol{s}}_{i-1}|\boldsymbol{s}_{i-1})P(\boldsymbol{s}_{i-1})P(\boldsymbol{s}_{i-2}, ..., s_{i-L})\\
=& \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1},\bar{\boldsymbol{s}}_{i-1}} P(\bar{\boldsymbol{s}}_{i-1}|\boldsymbol{s}_{i-1})(1 - Q(\lambda|_{\boldsymbol{s}_{i-1}, s_i=1}, \lceil \tau|_{\bar{\boldsymbol{s}}_{i-1}} \rceil))
\end{aligned} \quad (32)$$

The proof follows.

B. PROOF OF THE BER OF THE MULTI-BIT MEMORY RECEIVER

From (30), we have the following:

$$\begin{aligned}
&P(\bar{s}_i = 1|s_i = 0)\\
=& \sum_{\boldsymbol{s}_{i-1},\bar{s}_{i-1},...,\bar{s}_{i-K}} P(\bar{s}_i = 1|s_i = 0, \boldsymbol{s}_{i-1}, \bar{s}_{i-1}, ...\bar{s}_{i-K})\\
&P(\boldsymbol{s}_{i-1}, \bar{s}_{i-1}, ..., \bar{s}_{i-K})\\
=& \sum_{\boldsymbol{s}_{i-1},\bar{s}_{i-1},...,\bar{s}_{i-K}} P(\bar{s}_i = 1|s_i = 0, \boldsymbol{s}_{i-1}, \bar{s}_{i-1}, ...\bar{s}_{i-K})\\
&P(\bar{s}_{i-1}|s_{i-1}, ..., s_{i-L}, \bar{s}_{i-2}, ..., \bar{s}_{i-K})P(s_{i-1})\\
&P(s_{i-2}..., s_{i-L}, \bar{s}_{i-2}, ..., \bar{s}_{i-K})\\
=& \sum_{\boldsymbol{s}_{i-1},\bar{s}_{i-1},...,\bar{s}_{i-K}} P(\bar{s}_i = 1|s_i = 0, \boldsymbol{s}_{i-1}, \bar{s}_{i-1}, ...\bar{s}_{i-K})\\
&P(\bar{s}_{i-1}|s_{i-1}, ..., s_{i-L}, \bar{s}_{i-2}, ..., \bar{s}_{i-K})P(s_{i-1})\\
&P(\bar{s}_{i-2}|s_{i-2}..., s_{i-L}, \bar{s}_{i-3}, ..., \bar{s}_{i-K})P(s_{i-2}) \cdots\\
&P(\bar{s}_{i-K}|s_{i-K}, ..., s_{i-L})P(s_{i-K})P(s_{i-K+1}, ..., s_{i-L})
\end{aligned} \quad (33)$$

The term $P(\bar{s}_{i-1}|s_{i-1}..., s_{i-L}, \bar{s}_{i-L})$ can be calculated as follows:

$$\begin{aligned}
&P(\bar{s}_{i-1}|s_{i-1}, ..., s_{i-L}, \bar{s}_{i-2}, ..., \bar{s}_{i-K})\\
=& \sum_{s_{i-L-1},\bar{s}_{i-K-1}} P(\bar{s}_{i-K-1}|s_{i-K-1})P(s_{i-L-1})\\
&P(\bar{s}_{i-1}|s_{i-1}, ..., s_{i-L-1}, \bar{s}_{i-2}, ..., \bar{s}_{i-K-1})
\end{aligned}$$

Also, $P(\bar{s}_{i-2}|s_{i-2}..., s_{i-L}, \bar{s}_{i-3}, ..., \bar{s}_{i-K})$ can be obtained as follows:

$$\begin{aligned}
&P(\bar{s}_{i-2}|s_{i-2}, ..., s_{i-L}, \bar{s}_{i-3}, ..., \bar{s}_{i-K})\\
=& \sum_{s_{i-L-1},s_{i-L-2},\bar{s}_{i-K-1},\bar{s}_{i-K-2}} P(\bar{s}_{i-K-2}|s_{i-K-2})\\
&P(\bar{s}_{i-K-1}|s_{i-K-1}, s_{i-K-2}, \bar{s}_{i-K-2})P(s_{i-L-1})P(s_{i-L-2})\\
&P(\bar{s}_{i-2}|s_{i-2}, ..., s_{i-L-2}, \bar{s}_{i-3}, ..., \bar{s}_{i-K-2})
\end{aligned} \quad (34)$$

To obtain a tractable closed-form expression, we use the following approximation to compute, recursively, $P(\bar{s}_i = 1|s_i = 0)$:

$$\begin{aligned}
&P(\bar{s}_i = 1|s_i = 0)\\
=& \frac{1}{2} \sum_{s_{i-1},\bar{s}_{i-1}} P(\bar{s}_{i-1}|s_{i-1})P(\bar{s}_i = 1|s_i = 0, s_{i-1}, \bar{s}_{i-1})\\
\approx& \frac{1}{2^2} \sum_{s_{i-1},\bar{s}_{i-1}} P(\bar{s}_{i-1}|s_{i-1}) \sum_{s_{i-2},\bar{s}_{i-2}} P(\bar{s}_{i-2}|s_{i-2})\\
&P(\bar{s}_i = 1|s_i = 0, s_{i-1}, \bar{s}_{i-1}, s_{i-2}, \bar{s}_{i-2})\\
\approx& \frac{1}{2^L} \sum_{s_{i-1},\bar{s}_{i-1}} P(\bar{s}_{i-1}|s_{i-1}) ... \sum_{s_{i-K},\bar{s}_{i-K}} P(\bar{s}_{i-K}|s_{i-K})\\
&P(\bar{s}_i = 1|s_i = 0, \boldsymbol{s}_{i-1}, \bar{s}_{i-1}, ..., \bar{s}_{i-K})
\end{aligned} \quad (35)$$

The calculation of $P(\bar{s}_i = 1|s_i = 0, \boldsymbol{s}_i, \bar{s}_{i-1}, ..., \bar{s}_{i-K})$ can be done for any threshold, $\tau|_{\bar{s}_{i-1},...,\bar{s}_{i-K}}$, and from the

average number of received particles as follows:

$$\lambda_{\boldsymbol{s}_{i-1},s_i=0} \quad = \quad \lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j.$$

From (7), we obtain:

$$\begin{aligned} & P(\bar{s}_i = 1 | s_i = 0) \\ = & \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} \sum_{\bar{s}_{i-1},\ldots,\bar{s}_{i-K}} P(\bar{s}_{i-1}|s_{i-1}) \ldots P(\bar{s}_{i-K}|s_{i-K}) \\ & Q(\lambda_{\boldsymbol{s}_{i-1},s_i=0}, \lceil \tau |_{\bar{s}_{i-1},\ldots,\bar{s}_{i-K}} \rceil) \end{aligned} \quad (36)$$

$P(\bar{s}_i = 0 | s_i = 1)$ can be computed by using similar steps and assumptions:

$$\begin{aligned} & P(\bar{s}_i = 0 | s_i = 1) \\ \approx & \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} \sum_{\bar{s}_{i-1},\ldots,\bar{s}_{i-K}} P(\bar{s}_{i-1}|s_{i-1}) \ldots P(\bar{s}_{i-K}|s_{i-K}) \\ & P(\bar{s}_i = 0 | s_i = 1, \boldsymbol{s}_{i-1}, \bar{s}_{i-1}, \ldots, \bar{s}_{i-K}) \end{aligned} \quad (37)$$

where:

$$\lambda_{\boldsymbol{s}_i,s_i=1} \quad = \quad C_0 + \lambda_0 T + \sum_{j=1}^{L} s_{i-j} C_j$$

Finally, we obtain the following:

$$\begin{aligned} & P(\bar{s}_i = 0 | s_i = 1) \\ = & \frac{1}{2^L} \sum_{\boldsymbol{s}_{i-1}} \sum_{\bar{s}_{i-1},\ldots,\bar{s}_{i-K}} P(\bar{s}_{i-1}|s_{i-1}) \ldots P(\bar{s}_{i-K}|s_{i-K}) \\ & (1 - Q(\lambda_{\boldsymbol{s}_{i-1},s_i=1}, \lceil \tau |_{\bar{s}_{i-1},\ldots,\bar{s}_{i-K}} \rceil)) \end{aligned} \quad (38)$$

This concludes the proof.

REFERENCES

[1] S Hiyama, Y Moritani, T Suda, R Egashira, A Enomoto, M Moore, and T Nakano. Molecular Communication. 2005.

[2] Nariman Farsad, H Birkan Yilmaz, Andrew Eckford, Chan-Byoung Chae, and Weisi Guo. A comprehensive survey of recent advancements in molecular communication. IEEE Communications Surveys & Tutorials, 18(3):1887–1919, 2016.

[3] Massimiliano Pierobon and Ian F Akyildiz. Capacity of a diffusion-based molecular communication system with channel memory and molecular noise. IEEE Transactions on Information Theory, 59(2):942–954, 2013.

[4] S Mohammadreza Rouzegar and Umberto Spagnolini. Channel estimation for diffusive mimo molecular communications. In Networks and Communications (EuCNC), 2017 European Conference on, pages 1–5. IEEE, 2017.

[5] Vahid Jamali, Arman Ahmadzadeh, Christophe Jardin, Heinrich Sticht, and Robert Schober. Channel estimation for diffusive molecular communications. IEEE Transactions on Communications, 64(10):4238–4252, 2016.

[6] Ge Chang, Lin Lin, and Hao Yan. Adaptive detection and isi mitigation for mobile molecular communication. IEEE transactions on nanobioscience, 17(1):21–35, 2018.

[7] Yuting Fang, Adam Noel, Nan Yang, Andrew W Eckford, and Rodney A Kennedy. Symbol-by-symbol maximum likelihood detection for cooperative molecular communication. arXiv preprint arXiv:1801.02890, 2018.

[8] Vahid Jamali, Nariman Farsad, Robert Schober, and Andrea Goldsmith. Non-coherent detection for diffusive molecular communication systems. IEEE Transactions on Communications, 2018.

[9] Adam Noel, Karen C Cheung, and Robert Schober. Improving receiver performance of diffusive molecular communication with enzymes. IEEE Transactions on NanoBioscience, 13(1):31–43, 2014.

[10] Reza Mosayebi, Amin Gohari, Mahtab Mirmohseni, and Masoumeh Nasiri-Kenari. Type-based sign modulation and its application for isi mitigation in molecular communication. IEEE Transactions on Communications, 66(1):180–193, 2018.

[11] Po-Jen Shih, Chia-han Lee, and Ping-Cheng Yeh. Channel codes for mitigating intersymbol interference in diffusion-based molecular communications. In 2012 IEEE global communications conference (GLOBECOM), pages 4228–4232. IEEE, 2012.

[12] Deniz Kilinc and Ozgur B Akan. Receiver design for molecular communication. IEEE Journal on Selected Areas in Communications, 31(12):705–714, 2013.

[13] Martin Damrath and Peter Adam Hoeher. Low-complexity adaptive threshold detection for molecular communication. IEEE transactions on nanobioscience, 15(3):200–208, 2016.

[14] Reza Mosayebi, Hamidreza Arjmandi, Amin Gohari, Masoumeh Nasiri-Kenari, and Urbashi Mitra. Receivers for diffusion-based molecular communication: Exploiting memory and sampling rate. IEEE Journal on Selected Areas in Communications, 32(12):2368–2380, 2014.

[15] Hao Ye, Geoffrey Ye Li, and Biing-Hwang Juang. Power of deep learning for channel estimation and signal detection in ofdm systems. IEEE Wireless Communications Letters, 7(1):114–117, 2018.

[16] P. Dong, H. Zhang, and G. Y. Li. Machine learning prediction based csi acquisition for fdd massive mimo downlink. In 2018 IEEE Global Communications Conference (GLOBECOM), pages 1–6, Dec 2018.

[17] X. Qian, H. Deng, and H. He. Joint synchronization and channel estimation of aco-ofdm systems with simplified transceiver. IEEE Photonics Technology Letters, 30(4):383–386, Feb 2018.

[18] H Birkan Yilmaz, Changmin Lee, Yae Jee Cho, and Chan-Byoung Chae. A machine learning approach to model the received signal in molecular communications. In 2017 IEEE International Black Sea Conference

on Communications and Networking (BlackSeaCom), pages 1–5. IEEE, 2017.

[19] Changmin Lee, H Birkan Yilmaz, Chan-Byoung Chae, Nariman Farsad, and Andrea Goldsmith. Machine learning based channel modeling for molecular mimo communications. In 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pages 1–5. IEEE, 2017.

[20] Nariman Farsad and Andrea Goldsmith. Sliding bidirectional recurrent neural networks for sequence detection in communication systems. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018.

[21] X. Qian and M. Di Renzo. Receiver design in molecular communications: An approach based on artificial neural networks. In 2018 15th International Symposium on Wireless Communication Systems (ISWCS), pages 1–5, Aug 2018.

[22] Yansha Deng, Adam Noel, Weisi Guo, Arumugam Nallanathan, and Maged Elkashlan. Analyzing large-scale multiuser molecular communication via 3-d stochastic geometry. IEEE Transactions on Molecular, Biological and Multi-Scale Communications, 3(2):118–133, 2017.

[23] H Birkan Yilmaz, Akif Cem Heren, Tuna Tugcu, and Chan-Byoung Chae. Three-dimensional channel characteristics for molecular communications with an absorbing receiver. IEEE Communications Letters, 18(6):929–932, 2014.

[24] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. Deep learning, volume 1. MIT press Cambridge, 2016.

[25] Adam Noel, Karen C Cheung, and Robert Schober. Using dimensional analysis to assess scalability and accuracy in molecular communication. In 2013 IEEE International Conference on Communications Workshops (ICC), pages 818–823. IEEE, 2013.